AD-A053 145 JOHNS HOPKINS UNIV BALTIMORE MD F/G 12/1
CONSISTENCY OF A CLASS OF ROBUST ESTIMATORS OF CROSSCORRELATION--ETC(U)
MAR 78 V D VANDELINDE DAAG29-77-G-0093

UNCLASSIFIED ARO-14688.1-R-M NL

| OF |

AD
A053145

END
DATE
FILMED
6-78
DDC

ARO 14688.1-R-M



# CONSISTENCY OF A CLASS OF ROBUST ESTIMATORS OF CROSSCORRELATION

12

FINAL REPORT

V. David VandeLinde

March 31, 1978

U. S. Army Research Office

Grant Number: DAAG 29-77-G-0093

The Johns Hopkins University



Approved for Public Release;
Distribution unlimited.

| REPORT DOCUMENTATION PAGE | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|
| 1. REPORT NUMBER (18) ARO (19) 14688.1-R-M — 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle) (6) CONSISTENCY OF A CLASS OF ROBUST ESTIMATORS OF CROSSCORRELATION. | 5. TYPE OF REPORT & PERIOD COVERED (9) Final rep[illegible]t. |
| | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s) (10) V. David VandeLinde | 8. CONTRACT OR GRANT NUMBER(s) (15) DAAG 29-77-G-0093 new |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS The Johns Hopkins University | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (11) 31 Mar 78 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709 | 12. REPORT DATE March 31, 1978 |
| | 13. NUMBER OF PAGES 22 (12) 25p. |
| 14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) | 15. SECURITY CLASS. (of this report) Unclassified |
| | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE NA |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

NA

18. SUPPLEMENTARY NOTES

Prepared in cooperation with Santanu Basu

The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Robust estimation
Linear Systems
System Identification
Crosscorrelation estimators
Recursive algorithms

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In a previous work, briefly described in Section 1 of this report, a class of non-linear recursive estimators of cross-correlation was shown through simulations to be robust in identifying the parameters of linear, stationary, single-input-single-output systems whose output measurements are contaminated by noise which is not completely specified: the measurement noise distribution F is given by $F=(1-\varepsilon)K+\varepsilon C$, where K (over)

epsilon next page



191 750

Gue

ABSTRACT (cont.)

is completely known and C belongs to the class of zero-mean, symmetric, finite variance distributions.

The principal result of this report, described in Section 2, is a complement to these earlier results -- namely, a proof of consistency, with mean square convergence, for a general sub-class of the non-linear estimators of crosscorrelation defined in the earlier work. The proof is along the same lines as those followed in the Robbins-Monro stochastic approximation method.

# Table of Contents

# Section 1

# BACKGROUND AND PROBLEM STATEMENT

## 1.1 Introduction and background

The research performed under this grant (DAAG 29-77-G-0093) is a continuation of the work done by the author while participating in the Laboratory Research Cooperative Program (September 1975 - May 1976), Task Order 76-20, Basic Agreement DAHCO4 72-A-0001. The contents of this report should thus be considered as a complement to the final report (VandeLinde[1]) for the work just mentioned above. In the rest of this introduction, we give a brief description of the problem considered in [1]. Next, in 1.2, we state the results obtained therein, and finally, in 1.3, we describe the specific task undertaken in our research (i.e., under grant #DAAG 29-77-G-0093).

The following identification problem was considered in [1]: Let $\theta$ be a vector of unknown parameters $[a_1 \ldots a_p \; b_1 \ldots b_p]$ of a linear, single-input single-output system modeled in discrete time by

$$y(n)+a_1 y(n-1)+\ldots+a_p y(n-p)=b_1 u(n-1)+\ldots+b_p u(n-p) \qquad (1\ 1)$$

$$z(n)=y(n)+w(n) \qquad (1.2)$$

where the following conditions are assumed to hold:

Assumption 1 The linear system is stationary and the order $p$ is known;

Assumption 2 The system inputs $\{u(n)\}$ are free from observation errors and their statistical characteristics are completely known;

Assumption 3 The polynomials $A(x) = 1+a_1x+\ldots+a_px^p$ and $B(x) = b_1+b_1x+\ldots+b_px^p$ have no common factors;

Assumption 4 The inputs $\{u(n)\}$ and the measurement noise $\{w(n)\}$ are both zero-mean, finite-variance, ergodic sequences independent of each other;

Assumption 5 The probability distribution $F$ of $w(n)$ is given by

$$F = (1-\varepsilon)K+\varepsilon C \tag{1.3}$$

where $K\rightarrow$ a symmetric, zero-mean, finite variance distribution which is completely specified

$C\rightarrow$ a distribution which is only specified as belonging to $S_C$, the class of all zero-mean, finite variance, symmetric distributions

$\varepsilon\rightarrow$ a fixed number, $0\leq\varepsilon<1$, which may not be known. Typically, however, $\varepsilon<0.2$. As $C$ varies over $S_C$, (1.3) defines a convex set $P$ for possible $F$.

Using only the measured inputs $\{u(n)\}$ and measured outputs $\{z(n)\}$, the IDENTIFICATION TASK is to construct an on-line identification scheme whose performance -- as measured by the variance (or root-mean-square error) of estimation and speed of convergence -- is more or less uniform, i.e. robust, over different distributions for $w(n)$. Next, in 1.2, we briefly describe the results obtained in [1] for this (ROBUST) IDENTIFICATION TASK (R.I.T.).

In closing, we mention here that the estimation of $\theta$ under Assumptions 1-4 only has been the most widely studied problem in system identification (Åström and Eykhoff [2], Eykhoff [3] and Box and Jenkins [4]). In all such studies, though, the only concern has been to estimate the parameters with no consideration of desensitizing identification performance to the distribution of

$w(n)$ . Assumption 5 was added in the problem formulation in [1] for two reasons. Firstly, it was deemed reasonable to assume some, but not complete, knowledge of the operating environment. Additionally, since any parameter estimation scheme is based on the assumed model for $w(n)$ , such a model (Assumption 5), which allows for variations in the characteristics of $w(n)$ , is clearly necessary to carry out the task of constructing an identification procedure that works more or less uniformly well in a variety of situations.

## 1.2 Previous results

The approach in [1] for the stated R.I.T. is now described. A review of comparative evaluations of the existing identification methods (Isermann, Baur, Bamberger, Kneppo and Siebert [5], Saridis [6] and Sinha and Sen [7]) established that the method of correlation-- henceforth referred to as CR -- has the best performance as an identification scheme for the parameter estimation of systems of (1.1) and (1.2) under Assumption 1-4. Additionally and more importantly, while the $\{w(n)\}$ used in [5],[6] and [7] were quite different, CR was <u>uniformly</u> the best method in each study. Given these facts, the reasoning in [1] was straightforward -- with the addition of Assumption 5, CR could still be expected to have the best performance for <u>a particular</u> $F$ , although for different $F$ in $P$ , the performance of CR itself would vary. Thus, desensitizing the performance of CR to the distribution of $w(n)$ without incurring a loss of its existing advantages would provide a solution to the ROBUST IDENTIFICATION TASK. Before describing the modification of CR done in [1], we briefly recapitulate the essentials of the method.

With the assumption of ergodicity, we have the autocorrelation of the input given by

$$R_{uu}(m) = E[u(n)u(n-m)] = \lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} u(n)u(n-m) \tag{1.4}$$

and the crosscorrelation between the output and the input given by

$$R_{zu}(m) = E[z(n)u(n-m)] = \lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} z(n)u(n-m) \tag{1.5}$$

Since $E[w(n)u(m)]=0$ for all $n,m$, we also have

$$\begin{aligned} R_{zu}(m) &= E[z(n)u(n-m)] \\ &= E[y(n)u(n-m)]+E[w(n)u(n-m)] \\ &= E[y(n)u(n-m)] \\ &= R_{yu}(m) \end{aligned} \tag{1.6}$$

The convolution equation

$$R_{yu}(m) = \sum_{n=0}^{\infty} h_n \, R_{uu}(m-n) \tag{1.7}$$

relates $R_{yu}$ to $R_{uu}$ through the impulse response $\{h_k\}$ of the system. If the input autocorrelation is known, then (1.7) provides estimates of the impulse response from estimates $\hat{R}_{zu}(m)$ of the crosscorrelation since $R_{zu}(m)=R_{yu}(m)$ from (1.6). In particular, using a white $\{u(n)\}$, for which $R_{uu}(n)=0$ for all $n\neq 0$, the $\{\hat{h}_k\}$ are simply obtained from

$$\hat{h}_k = \frac{\hat{R}_{zu}(k)}{R_{uu}(0)} \tag{1.8}$$

The parameters $[a_1 \ldots . a_p \; b_1 \ldots . b_p]$ can then be estimated from the $\{\hat{h}_k\}$ by a least squares procedure -- see for instance, p.32, Section 1 of [1].

The modification of CR in [1] was a non-linear estimation of the crosscorrelation instead of the usual linear scheme. We give below the general non-linear form, of which the linear scheme is just a special case with $g(\cdot)$ as the identity function:

$$\hat{R}_{zu}(m,n)=\hat{R}_{zu}(m,n-1)-A(n)g[\hat{R}_{zu}(m,n-1)-a(n)u(n-m)] \tag{1.9}$$

with $n=0,1,2,...$

$$A(n)=1/n+1$$

$$\hat{R}_{zu}(m,-1)=0$$

and

$$\begin{aligned} g(x) &= s_1 x \quad |x|<d_1 \\ &= s_1 d_1 \operatorname{sgn} x \quad d_1 \le |x|<d_2 \\ &= s_2(x - H \operatorname{sgn} x) \quad d_2 \le |x| < H \\ &= 0 \quad |x| \ge H \end{aligned} \tag{1.10}$$

$$H = \left(\frac{s_2 d_2 - s_1 d_1}{s_2}\right)$$

$$s_1>0, s_2 \le 0, d_2>d_1>0 \ .$$

$(s_1,s_2,d_1,d_2)$ were chosen as functions of the parameters of $K$, the specified part of $F$. As with the unmodified CR, $[a_1....a_p \ b_1....b_p]$ were then estimated from the $\{\hat{R}_{zu}(m)\}$.

Extensive computer simulations were carried out in [1], with 20 different distributions for $w(n)$, seven sets of values for the parameters $(s_1,s_2,d_1,d_2)$ and three linear processes to study the performance of (the modified) method of correlation for different $w(n)$ using these non-linear estimators of $R_{zu}(m)$. One of the seven sets of values of $(s_1,s_2,d_1,d_2)$ corresponded to the linear estimator (i.e. $g(x)=x$ for all $x$), which is the common method of estimating crosscorrelations on-line. The result of [1] showed quite clearly that identification of $[a_1....a_p \ b_1....b_p]$ using the non-linear estimates of $R_{zu}$ was indeed 'robust' in the sense of better and more uniform performance than CR for such

varying $w(n)$ . However, some important work remained to be done. We next describe the specific task which arises naturally from the work of [1] and which we have undertaken in our research.

## 1.3 Problem statement

There is a primary need, to show for any estimation scheme, that the procedure is consistent in estimating the unknown parameters. Although the simulations in [1] did not yield any divergent results, the consistency of the non-linear estimators of $R_{zu}$ (given by (1.9)) were not analytically shown. This is the specific task undertaken in this research: to (analytically) demonstrate the consistency of the 'robust' estimators of $R_{zu}$ as defined by equation (1.9).

# Section 2

## CONSISTENCY OF SOME ROBUST ESTIMATORS OF CROSSCORRELATION

### 2.1 Notation, assumptions and preliminary results

For notational convenience, we re-write equation (1.9) as

$$x(n)=x(n-1)-A(n)g[x(n-1)-z(n)u(n-m)] \tag{2.1}$$

with $n=0,1,2,...$

$x(-1)=0$

$A(n)=1/n+1$

and $g(\cdot)$ is defined by (1.10)

Clearly, $x(n)$ represents $\hat{R}_{zu}(m,n)$ . Note that $m$ can be supressed in the index of $x$ , since it is fixed for a set of iterations.

To prove the consistency of (2.1), we need two additional conditions (besides Assumptions 1-5):

Assumption 6 $\{u(n)\}$ is an independent, binary sequence;

Assumption 7 $g(\cdot)$ is the so-called 'soft-limiter' function, defined as

$$\begin{aligned} g(x) &= x \quad |x| < d_1 \\ &= d_1 \operatorname{sgn} x \quad |x| \geq d_1 \end{aligned} \tag{2.2}$$

Clearly, equation (2.2) is a special case of equation (1.10) with $s_1=1, s_2=0, d_2=\infty$ . As before, $d_1$ is chosen as a function of parameters of $K$ , the known part of the distribution for $w(n)$ .

Remark: Since the method of correlation is most often implemented during normal operation with an independent, binary sequence as the common choice of test signal (Sage and Melsa [8]), Assumption 6 is not an unrealistic condition.

Our method of proof needs, not surprisingly, results regarding properties of the probability distribution of $z(n)u(n-m)$ for systems of (1.1) and (1.2) under Assumptions 1-7. We consider first the more general case with Assumption 6 replaced by Assumption 6-a:

Assumption 6-a $\{u(n)\}$ is an identically distributed, independent sequence with a zero mean symmetric distribution (not necessarily binary).

Then for zero initial conditions and $u(n)=0$ for $n<0$, we may write

$$y(n) = \sum_{j=0}^{n-1} h_{n-j} u(j) \tag{2.3}$$

for the natural, undisturbed output of a system defined by equation (1.1). Thus,

$$z(n)u(n-m) = \left[ \sum_{j=0}^{n-1} h_{n-j} u(j) + w(n) \right] u(n-m)$$

$$= \underbrace{\sum_{\substack{j=0 \\ j\neq(n-m)}}^{n-1} h_{n-j} u(j) u(n-m)}_{\text{Term I}} + \underbrace{w(n)u(n-m)}_{\text{Term II}} + \underbrace{h_m u^2(n-m)}_{\text{Term III}} \tag{2.4}$$

We now state our first result:

Lemma 1: If $\{u(n)\}$ is independent, zero-mean and symmetrically distributed, the distribution of $z(n)u(n-m)$ given $u(n-m)$ is symmetric around its mean of $h_m u^2(n-m)$.

We write $[A|B]$ to denote a random variable $A$ given a random variable $B$.

Proof: The distribution of $[z(n)u(n-m)|u(n-m)]$ is simply the distribution of $z(n)$ scaled by a constant $(= u(n-m))$. So we consider equation (2.4) with $u(n-m)$ now a constant.

Term I now is just a sum of zero-mean, independent random variables with (identical) symmetric distributions, and so is itself a random variable with a zero-mean and a symmetric distribution (Papoulis [9]). Term II is also a zero-mean, symmetrically distributed random variable, since $w(n)$ is such and the mean and symmetry (of a random variable) are unaffected under multiplication by a constant. Also, Term II is independent of Term I.

Term III, however, is a constant, since $u(n-m)$ is given, and not a random variable.

The stated result (Lemma 1) immediately follows.

Remark: Lemma 1 states that

$$E[z(n)u(n-m)|u(n-m)]=h_m u^2(n-m) \tag{2.5}$$

where $E[\cdot]$ denotes the expectation operator. Remembering that $E[A]=E_B[E[A|B]]$, where the outer expectation $E_B$ is over $B$, we may check our result (of Lemma 1) by computing $E[z(n)u(n-m)]$ from $E_u[E[z(n)u(n-m)|u(n-m)]]$. Thus,

$$\begin{aligned} E[z(n)u(n-m)] &= E_u[E[z(n)u(n-m)|u(n-m)]] \\ &= E_u[h_m u^2(n-m)] \quad \text{(from Lemma 1)} \\ &= h_m E_u[u^2(n-m)] \\ &= h_m R_{uu}(0) \end{aligned} \tag{2.6}$$

We have already seen that $R_{zu}(m)=E[z(n)u(n-m)] = h_m R_{uu}(0)$ for white (zero-mean, independent) $\{u(n)\}$, so our result checks. Henceforth, we refer to $h_m R_{uu}(0)$ by $R$.

If we impose the more restrictive Assumption 6 instead of 6-a, that is, $\{u(n)\}$ is now an independent, binary sequence, the following result is obtained:

Lemma 2: If $\{u(n)\}$ is an independent, binary sequence, $z(n)u(n-m)$ has a symmetric distribution around a mean $R(=h_m R_{uu}(0))$ .

Proof: From Lemma 1, the distribution of $[z(n)u(n-m)|u(n-m)]$ is symmetric around its mean $h_m u^2(n-m)$ .

Since $\{u(n)\}$ is binary, $u^2(n)$=constant for all $n$ . In other words,

$$\begin{aligned} R_{uu}(0) &= E[u^2(n)] \\ &= \text{constant} \qquad\qquad (2.7) \\ &= u^2(n-m) \text{ , in particular} \end{aligned}$$

That is, the distribution of $[z(n)u(n-m)|u(n-m)]$ is symmetric around a mean $h_m R_{uu}(0)(=R)$ which is independent of $u(n-m)$ , the conditioning random variable. Thus, our result follows.

Remark: We could as well prove Lemma 2 by directly considering equation (2.4).

Term I is the sum of products of pairs of independent, identically distributed random variables which are zero-mean and symmetric and so is itself zero-mean, symmetrically distributed.

Similarly, Term II is also a zero-mean, symmetrically distributed random variable. The sum of Terms I and II is still zero mean, symmetrically distributed. Since Term III, $h_m u^2(n-m)$ , is a constant for binary $\{u(n)\}$ , Lemma 2 follows.

We require one more preliminary result before proceeding to the proof of consistency for 2.1:

Lemma 3: For $\{u(n)\}$ independent and binary,

$$E[g[R-z(n)u(n-m)]] = 0 \tag{2.8}$$

Proof: The distribution of $z(n)u(n-m)$ is symmetric around its mean $R$ (Lemma 2). Define a new random variable $Q = R-z(n)u(n-m)$. Clearly, $Q$ has a symmetric distribution about a mean of zero. Now,

$$E[g[R-z(n)u(n-m)]]$$
$$= \int g[R-z(n)u(n-m)]\, dD_Z \tag{2.9}$$

where $D_Z$ is the distribution of $z(n)u(n-m)$. Substituting $Q$ for $[R-z(n)u(n-m)]$, we can write

$$E[g[R-z(n)u(n-m)]]$$

$$= \int g(Q)\, dD_Q \tag{2.10}$$

where $D_Q$ is the distribution of $Q$.

Since $g(\cdot)$ is an odd, bounded continuous function, our desired result follows.

We are now in a position to state and prove our main result, which we do in 2.2.

## 2.2 Proof of Consistency

The principal result of our work may be stated as:

Theorem

Given a linear system described by equations (1.1) and (1.2) under Assumptions 1-6, equation (2.1) (or (1.9)) defines a crosscorrelation estimate that is consistent, with convergence in the mean square, for $g(\cdot)$ defined by Assumption 7.

In other words, under the assumed conditions,

$$\hat{R}_{zu}(m,n) \xrightarrow{m.s.} R_{zu}(m)$$

where $\xrightarrow{m.s.}$ denotes convergence in mean square.

<u>Proof</u>: We essentially follow the method of Robbins and Monro [10].

Recalling equation (2.1),

$$x(n) = x(n-1) - A(n)g[x(n-1)-z(n)\ u(n-m)] \tag{2.11}$$

$n = 0,1,2,...$

$x(-1) = 0$

$$A(n) = \frac{1}{n+1} \quad \sum_{n=1}^{\infty} A(n) = \infty \ ; \quad \sum_{n=1}^{\infty} A^2(n) < \infty$$

Remembering that $R_{zu}(m) = h_m\ R_{uu}(0) = R$ , we set

$$x(n) - R = [x(n-1) - R] - A(n)g[x(n-1)-z(n)u(n-m)] \tag{2.12}$$

Squaring both sides,

$$[x(n)-R]^2 = [x(n-1)-R]^2 + A^2(n)\{g[x(n-1)-z(n)u(n-m]\}^2$$
$$-2\ A(n)[x(n-1)-R]g[x(n-1)-z(n)u(n-m)] \tag{2.13}$$

Taking expectations w.r.t. $x(n-1)$,

$$E[\{x(n)-R\}^2|x(n-1)]=[x(n-1)-R]^2$$
$$+ A^2(n)E[\{g[x(n-1)-z(n)u(n-m)]\}^2\ |x(n-1)]$$
$$-2\ A(n)[x(n-1)-R]E[g[x(n-1)-z(n)u(n-m)]|x(n-1)] \tag{2.14}$$

Set $E[g[x(n-1) - z(n)u(n-m)]|\ x(n-1)] = \beta[x(n-1)]$

There are three cases to be considered:

(a) <u>$x(n-1)=R$</u>: $\beta[x(n-1)] = 0$ by Lemma 3. (2.15)

(b) $\underline{x(n-1)>R}$: Point by point $[x(n-1)-z(n)u(n-m)]>[R-z(n)u(n-m)]$, and since $g(\cdot)$ is an odd, bounded, continuous, monotonically increasing function,

$$\beta[x(n-1) > 0 \ . \tag{2.16}$$

(c) $\underline{x(n-1)<R}$: Point by point $[x(n-1)-z(n)u(n-m)]<[R-z(n)u(n-m)]$, and for reasoning similar to that in (b) above,

$$\beta[x(n-1)] < 0 \ . \tag{2.17}$$

Setting $E[x(n)-R]^2 = B(n)$ , and using (2.14), we obtain

$$B(n)=B(n-1)+A^2(n)E[E[\{g[x(n-1)-z(n)u(n-m)]\}^2|x(n-1)]]$$
$$-2\,A(n)E[\{x(n-1)-R\}\beta[x(n-1)]] \tag{2.18}$$

Set $E[E[\{g[x(n-1)-z(n)u(n-m)]\}^2|\,x(n-1)]] = f(n)$
and $E[\{x(n-1)-R\}\beta[x(n-1)]] = e(n)$ .

Since we are only considering distributions of finite variance, and $g(\cdot)$ is odd, continuous and bounded, it follows [10] that $0\le f(n)<\infty$ . Also, because of (2.15), (2.16) and (2.17),

$$e(n) \geq 0 \ .$$

We may re-write (2.18) as

$$B(n)=B(n-1) + A^2(n)f(n)-2\,A(n)e(n) \tag{2.19}$$

Summing over (2.19), we get

$$B(n) = B(1) + \sum_{j=1}^{n-1} A^2(j)f(j) - 2\sum_{j=1}^{n-1} A(j)\,e(j) \tag{2.20}$$

Note first that with $f(n) \geq 0$ , and $\sum_{j=1}^{\infty} A^2(j) < \infty$ , we have $\sum_{j=1}^{\infty} A^2(j)f(j) < \infty$ from [10].

Since $B(n) \geq 0$ , from (2.20) we have

$$\sum_{j=1}^{n-1} A(j)\, e(j) \leq 1/2\, [B(1) + \sum_{j=1}^{\infty} A^2(j) f(j)] < \infty \qquad (2.21)$$

Hence, the positive term series $\sum_{j=1}^{\infty} A(j)e(j)$ converges.

Thus, $$\lim_{n\to\infty} B(n) = B(1) + \sum_{j=1}^{\infty} A^2(j)f(j) - 2 \sum_{j=1}^{\infty} A(j)e(j)$$

$$= B \text{ exists and } B \geq 0 .$$

Since $A(n) = 1/n+1$ , we have from [10],

$$\lim_{n\to\infty} B(n)=B=0 .$$

i.e. $\lim_{n\to\infty} E[x(n)-R]^2 = 0$

or, $\hat{R}_{zu}(m,n) \to R_{zu}(m)$ in mean square .

<u>Remark 1</u>: For consistency, we only have to show that

$\lim_{n\to\infty} P[|\hat{R}_{zu}(m,n)-R_{zu}(m)|>\varepsilon] = 0$ , i.e. convergence in probability.

Our theorem demonstrates <u>mean square</u> convergence of $\hat{R}_{zu}(m,n)$ to $R_{zu}(m)$, and since convergence in mean square implies convergence in probability, we have established a stronger form of convergence than is required for consistency.

<u>Remark 2</u>: Some of the results of this section have been reported earlier (Basu and VandeLinde [11]), and details of all results (of Section 2) can be found in Basu [12].

## Section 3

## CONCLUDING REMARKS

The principal result of this research has been to show the consistency, with convergence in mean square, of the robust estimator of crosscorrelation described in [1]. There are several important issues, however, that remain to be addressed.

Clearly, the first obvious undertaking would be relaxations of Assumptions 6 and 7.

Specifically, one might attempt to prove consistency under the following sets of assumptions:

(a) Assumptions 1-5, 6-a and 7

(b) Assumptions 1 through 7, with 6-a replacing 6

(c) Assumptions 1 through 6

(d) Assumptions 1 through 5 and 6-a

(e) Assumptions 1 through 5 alone

We have listed the above sets (a)-(e) in increasing orders of generality (of problem formulation) <u>and</u> of difficulty (of solution). We should mention, though, that none of these tasks are trivial. In fact, our conjecture is that, for these possible relaxations of Assumptions 6 and 7, we might have to pursue a different approach to prove consistency. Recently, in 1977, Ljung [13] and Kushner [14] have proved convergence of certain classes of recursive stochastic algorithms using ideas of stability and weak convergence theory, respectively. While their results are not directly applicable to our case, their methods might still be of some benefit -- at the very least, they provide two alternative avenues of approach.

In a more general regard, it would be desirable to derive asymptotic properties of the recursive estimator of $R_{zu}$ -- namely, does the distribution of $\{\hat{R}_{zu}(m,n)\}$ converge to some stable distribution, hopefully normal? Asymptotic normality would provide the considerable benefit of an explicit expression for the asymptotic variance of the estimator, enabling in turn a more analytical treatment of the robust identification problem (see, for example, Appendix B of [1] for general ideas of precisely formulating a robust problem). Our recursive scheme, unfortunately, has an inherent lack of tractable mathematical structure for tackling such questions. We note that equation (2.1) defines a sequence $\{x(n)\}$ which is <u>not</u> stationary, <u>not</u> Markovian, nor does it have any martingale properties. It is small consolation to add that all general, stochastic, recursive algorithms operating on <u>dependent</u> data suffer from similar drawbacks and to date, no results have been obtained with regard to asymptotic distributions of such schemes.

Other questions of interest are choice of $d_1$ (for assumption 7) -- and choice of $(S_1,S_2,d_1,d_2)$ in the more general case -- for improving rate of convergence. Lastly, but not least, is the question of the 'best' transformation to use to obtain the estimates $\{\hat{a}_i,b_i\}$ from the $\{\hat{h}_i\}$ . For noisy $\{\hat{h}_i\}$ , as is obviously the case here, this is quite an open question.

## BIBLIOGRAPHY


[1] V. D. VandeLinde (1977), "Robust identification of linear systems", BRL Report No. 1960, U.S. Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland 21205.

[2] K. J. Åström and P. Eykhoff (1971), "System identification -- a survey", <u>Automatica</u>, Vol. 7, 123-162.

[3] P. Eykhoff (1974), "System identification -- parameter and state estimation", John Wiley & Sons, New York.

[4] G. E. P. Box and G. M. Jenkins (1976), "Time series analysis: forecasting and control", Holden-Day, San Francisco.

[5] R. Isermann, U. Baur, W. Bamberger, P. Kneppo and H. Siebert (1974), "Comparison of six on-line identification and parameter estimation methods", <u>Automatica</u>, Vol. 10, 81-103.

[6] G. N. Saridis (1974), "Comparison of six on-line identification algorithms", <u>Automatica</u>, Vol. 10, 69-79.

[7] N. K. Sinha and A. K. Sen (1975), "Critical evaluation of on-line identification methods", <u>Proc. IEE</u>, Vol. 122, 1153-1158.

[8] A. P. Sage and J. L. Melsa (1971), "System identification", Academic Press, New York.

[9] A. Papoulis (1967), "Probability, random variables and stochastic processes", McGraw-Hill, New York.

[10] H. Robbins and S. Monro (1951), "A stochastic approximation method", <u>Annals Math. Stat.</u>, Vol. 22, 400-407.

[11] S. Basu and V. D. VandeLinde (1977), "Robust identification of parameters of linear systems", Proceedings of the 15th Annual Allerton Conference on Communication, Control and Computing, Monticello, Illinois.

[12] S. Basu (1978), "Robust estimation of crosscorrelation for parameter identification of linear systems", unpublished Ph.D. dissertation, The Johns Hopkins University, 1978.

[13] L. Ljung (1977), "Analysis of recursive stochastic algorithms", IEEE Trans. Auto. Control, Vol. AC-22, No. 4, 551-575.

[14] H. J. Kushner (1977), "Convergence of recursive adaptive and identification procedures via weak convergence theory", IEEE Trans. Auto. Control, Vol. AC-22, No. 6, 921-930.

# Appendix A

## List of Publications

1. "Robust identification of parameters of linear systems", Santanu Basu and V. David VandeLinde, Proceedings of the 15th Annual Allerton Conference on Communication, Control and Computing, Monticello, Illinois, 1977.

# Appendix B

## List of Scientific Personnel Supported and Degrees Awarded

1. Santanu Basu: graduate student and candidate for Ph.D. degree (to be awarded May 1978).